

Emotion Model Implementation for Parameterized Facial Animation in Human-Robot-Interaction

Steffen Wittig^{1*}, Uwe Kloos¹, Matthias Rätsch²

¹ Computer Graphics, Reutlingen University, Reutlingen, Germany.

² Mechatronics, Reutlingen University, Reutlingen, Germany.

* Corresponding author. Tel: +49 7473 956 85 45; email: info@steffenwittig.com

Manuscript submitted July 9, 2015; accepted December 11, 2015.

doi: 10.17706/jcp.11.6.439-445

Abstract: In recent years robotic systems have matured enough to perform simple home or office tasks, guide visitors in environments such as museums or stores and aid people in their daily life. To make the interaction with service and even industrial robots as fast and intuitive as possible, researchers strive to create transparent interfaces close to human-human interaction. As facial expressions play a central role in human-human communication, robot faces were implemented with varying degrees of human-likeness and expressiveness. We propose an emotion model to parameterize a screen based facial animation via inter-process communication. A software will animate transitions and add additional animations to make a digital face appear “alive” and equip a robotic system with a virtual face. The result will be an inviting appearance to motivate potential users to seek interaction with the robot.

Key words: Facial animation, circumplex model, face design.

1. Introduction

Conventional robots work isolated from humans on repetitive tasks such as painting, welding and simple assembly tasks [1] but in recent years so called *socially assistive* or *socially interactive* robots have been developed to aid people in their daily life. They are able to perform home and office automation, work in cooperation with humans [2] and guide them in environments such as museums or stores. They can also be used as research platforms, toys, educational tools and therapeutic aids [3]. Socially interactive robots may be a key factor in battling serious problems that can be expected with our coming of age society, e.g. by performing simple health care and welfare related tasks in cooperation with humans [1].

Human-Robot Interaction (HRI) researchers suggest that humans prefer to interact with machines, like they interact with fellow humans [3]. To make the user feel empowered and competent, it is necessary to adhere to human social expectations to provide a natural, intuitive and enjoyable interface that needs no training to use it. Facial expressions are a fundamental aspect of social interactions and can be rapidly recognized, enabling us to draw certain conclusion about the inner state of fellow humans faster. Using facial expressions it is possible to translate certain states of the robotic system to facial expressions and add to the information provided by the already widely used text and voice output.

Various fields of study from psychologists to computer scientists work on synthesizing facial expressions [4] and several research groups have developed humanoid robots with animated faces. An example of this is Albert HUBO, a human scale humanoid robot with an almost lifelike face. Based on the biped walking robot HUBO, researchers have built a mechanical face and added a skin made out of a sponge-like elastomer

on top of it to create movement as close as possible to real facial soft-tissue [5]. The realistic human-like face can influence user acceptance negatively (see “Uncanny Valley” problem in section 2) and because of their precisely engineered mechanical parts, they may need considerably more work if the need to change their design arises and are likely to increase the need for maintenance.

Computer graphics based approaches can be realized by having a software solution continuously render an animated face and display this output on a display screen or project it into a flat, curved or even head-shaped surface, preferably via a back projection. Baxter is an industrial robot by Rethink Robotics, which is able to safely operate around human workers and communicate with them through a simplified face displayed on a display screen. Baxter can be trained to perform repetitive tasks by having a human trainer guide his arm to show it the needed motions and objects to perform a task. It is then able to repeat the task without any programming – Baxter can therefore be trained by users without software, robotics or engineering knowledge. A pair of eyes and eyebrows are displayed on Baxter's head mounted display to signal its human users whether Baxter was able to interpret the user's input through simplified expressions, as well as show where its arm will be going next by looking in that direction. Baxter's expression changes are quickly animated and limited to the angle of the eyebrows and eyelids, opening closing of the eyes and moving the pupils. This suffices for Baxter's use case and how facial expressions enhance the human-robot interfaces, even for industrial robots [6].

SociBot by Engineering Arts Limited features an animated face projected onto the roughly human-face-shaped head of a 60 centimeter high bust. The robot focuses on social interaction and allows to freely customize the displayed face's texture. Its facial expressions can be continuously transformed by manipulating values based on the extensive Facial Action Coding System (FACS), which is derived from the various muscles of a human face, by means of a custom-built facial animation software. Transitions between expressions are animated but the facial features remain static otherwise, except for blinking and talking animations. The makers of SociBot suggest that it could be used as a terminal in a mall, airport or bank. The robot can mechanically move its head to look at a person, in addition to the digital face's eye movement [7].

“Mask-Bot” is a similar system with a human-face-shaped mask mounted on a pan-tilt unit for mechanical head movement. It also allows to display freely customized and animated faces through back projection, creating a life-sized talking head in 3D space, as opposed to 2D flat surfaces of display screens. Its software allows for Text-to-Speech audio output to be accompanied by synchronous lip movements, based on phonemes. The easy to change appearance allows to rapidly prototype robot head designs and evaluate how they appear to users of robotic systems, addressing the problem of finding ideal face models and behaviors for platforms with mechanical faces, that are harder to modify. [8] Mask-Bot and SociBot enter Uncanny Valley territory once realistically textured faces are projected onto their human-like head shape, due to the noticeably mechanical head movements and animations.

1.1. Goal

To enhance the interaction with socially interactive systems the focus was put on creating a facial animation software with an underlying emotional model based on intuitive parameters, which can be sent by other processes. Liveliness animations, like breathing and blinking, will show users that the system is ready to be engaged in conversation and motivate them to interact with it. The parameters of the facial expressions are based on a two-dimensional expression model rather than on the parameter rich and abstract FACS. This paves the way for a meaningful and easy interface of the facial animation software. The face will be rendered on a screen and must be animated in a way that it appears alive at all times, as a visual cue that the host system is active.

It was hypothesized that facial expressions will support the robot's speech synthesis by allowing faster

interpretation of messages and also make the robot more sympathetic by showing an emotion based on its internal state, e.g. show a sad facial expression as soon as the robot starts to say the words “I could not understand your question.” This way the user understands that the machine empathizes with its human partner, who has to repeat what he/she just said.

2. Face Design

The face serves as the focal point for a human interaction partner – intuitive interaction will be encouraged by displaying clear communicative features. Facial features can be designed with visual cues that help users understand a robot's capabilities [9]. Studies suggest that human-like face designs can create expectations of human-like mind and abilities [10]. If the robot's conversational abilities, more precisely the speed of speech recognition and synthesis, are not on par with an adult human's abilities, it is therefore not advisable to design a robot's face realistically human-like.

The design should rather fit the tasks that the robot will perform and its abilities. Realistic depictions of human faces also tend to have imperfections that put them in the much discussed “Uncanny Valley” [11], which could result in feeling of eeriness for the users.

Our face design has to be able to display emotions that aid its spoken messages. Not every emotion is needed for the intended scenario. The robot's only purpose is to serve humans, so why should it display an angry face, for example? For a service based scenario, the following expressions were identified: *Happy*, for positive feedback; *Attentive*, for when the robot is listening to user input; *Sad*, for negative feedback.

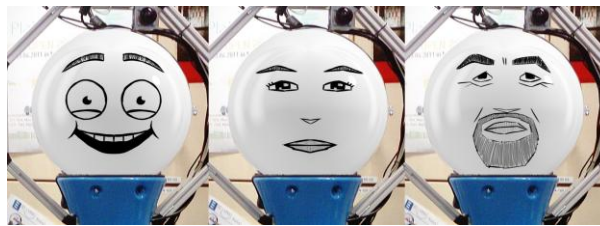


Fig. 1. From left to right: “Cartoon-like”, “female” and “male” face design sketches.

For the proposed work, it was decided to create three high-contrast face designs with few details: A more natural female and male face, as well as an abstract cartoon-like face (see Fig. 1). To evaluate what the best level of human-likeness is, without getting into “Uncanny Valley territory”, a questionnaire was designed to find out how sympathetic, trustworthy and natural the faces are perceived by humans. Also of interest was, how well specific emotions could be transported with the different levels of abstract faces.

For every face design neutral, happy, sad, attentive and frustrated expressions were sketched out and rendered on a photograph of a SCITOS G5 robotic system. The needed deformations to create these facial expressions for these emotion words are drawn from intuition and whether they are appropriate will be evaluated by the questionnaire. Participants were asked to distribute points to specific attributes, like “happy” or “inquiring”, for every expression of every face design.

The questionnaire was handed out to students of a computer graphics course, a computer science course and visitors of an open house event at Reutlingen University. 105 questionnaires were filled out with 84.8% of the participants being male. Participants were in the age range of 16 to 59 years. Because most questionnaires were filled out by Bachelor's degree students, most participants were between 18 and 24 years old.

The questionnaire results of female and male participants were weighted to compensate for more than 80% of them being male (see Fig. 2). Results indicate that the cartoon-like face design wasn't perceived as natural and trustworthy as the more human-like face designs. All of the male face design's attributes are

rated slightly higher than the female face design's attributes. In the end the female face design is chosen to be implemented in the prototype. It wasn't rated considerably less sympathetic, trustworthy and natural than the male face design but is more fitting for the use case and may work better in combination with the skirt-like body shape of the robot.

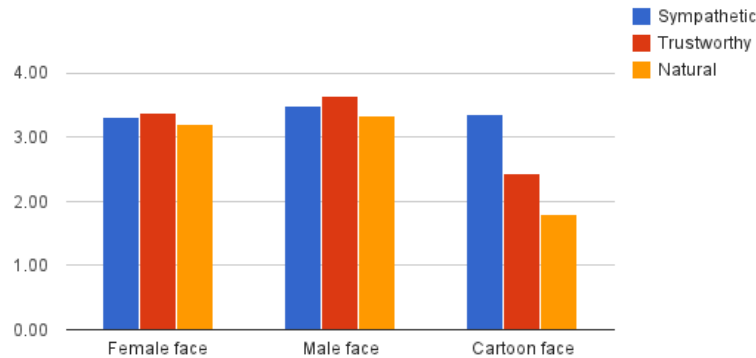


Fig. 2. Weighted female and male results of the questionnaire.

The male face design's expressions also seem to be harder to interpret, though this might be an issue with specific facial features of the male design (wrinkles, beard). The cartoon-like face design was apparently far less trustworthy than the human-like face designs and is therefore not a good choice without considerable changes to the design. Its expressions transport attributes similarly well as the female-face design. The female face design was then recreated as a 3D model and prepared for animation.

3. Implementation

3.1. Emotion Model

While a model like FACS allows to describe the change in face muscles needed to form emotive expressions, researchers have also developed models to describe human emotions via a set of parameters. One of those models is James A. Russell's "circumplex model of emotion", which was adapted for an HRI use case.

Russell took the everyday emotion categories (emotion words) and placed them in the form of a circle in a two-dimensional space. Besides the horizontal and vertical axis, the circular placement of the emotion words is what defined Russell's circumplex model of emotion by adding an additional layer to it: The words act like the hues in a chromatic circle and can blend into each other. Terms that are closely related will be closer together (happy and delighted) while antonyms (happy and sad) will be on the opposite of the circle. This allowed to specify a degree to which emotion categories are interrelated: The correlation equals the cosine of the angle between any two categories. This means that two emotion words with a cosine of 90° have nothing in common ($\cos 90^\circ = 0$) and words with a cosine of 180° describe exactly the opposite emotion ($\cos 180^\circ = -1.00$). In the center of Russell's model is a neutral point and the intensity of each emotion can be described by the length of a vector originating from it [12].

3.2. Software

The sketched out facial expressions were recreated and placed in a two-dimensional blend tree in the Unity engine, at the same positions as the corresponding emotion words in Russell's model (see Fig. 3).

Our software will react to messages sent via UDP, containing the keys *arousal* or *pleasure* and a value between -100 and 100. Upon receiving the message, an animation controller will set the received parameters and blend between the closest facial expressions accordingly and smoothly transition the

parameters via linear interpolation over a certain amount of time, which influences the expression of the 3D face model.

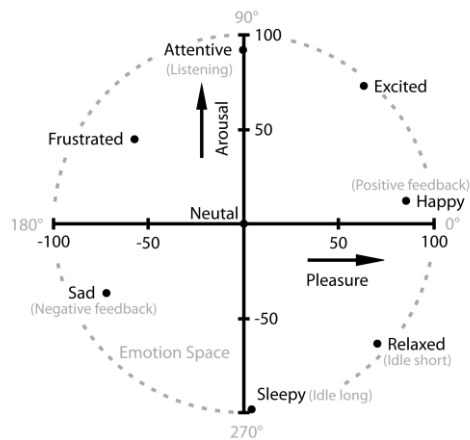


Fig. 3. The adapted circumplex model used for parameterizing facial expressions.

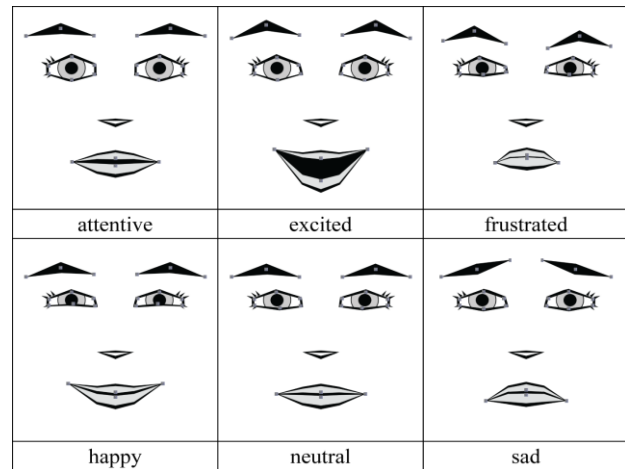


Fig. 4. Expressions realized in the proposed system.



Fig. 5. Previsualization. Output of our software rendered on top of a photograph of the target system.

Several animations (see Fig. 4) were created and added to the facial expressions to create a lively appearance: breathing, blinking, small random head movements and micro movements, like the slight twitching of an eyebrow. Besides “arousal” and “pleasure”, 3D gaze coordinates, “talking” and “idle” switches can be sent via an UDP interface, to have the eyes and head look at specified coordinates, to activate or deactivate a generic talking animation or enter an idle mode, which will randomly select emotion and gaze parameters from time to time.

To evaluate the facial animation software a Java program was implemented, which allows to send messages in addition to playing prerecorded audio clips for a service related scenario. The scenario would have the robot (see Fig. 5) approaching a customer in a hardware store and greet them (happy expression), ask what item they are looking for and waits for input (attentive expression). If it could not understand the customer or the item is not in stock, it gives negative feedback (sad face), otherwise it gives positive feedback (happy face) and asks the user to follow it. Upon reaching the location of the item the robot would show the user through gaze, where the item can be found.

The scenario was performed on a monitor in front of 12 participants. They were asked to rate how they

perceived attributes of the character via a short questionnaire: It was perceived as mostly sympathetic (8.18/10), trustworthy (8.0/10) and intelligible (9.18/10), but not as believable (6.77/10), which some of them noted, was due to noticeably asynchronous lip movements hurting the illusion of a live being.

4. Conclusion

Through empiric research and a stochastic approach, a balance was found for the design to enhance human-robot-interaction. The parameters exposed through the software's interface relate to meaningful position of emotions in a two-dimensional emotion space, in contrast to a rich system like FACS, that is abstract and not intuitive to use. However, using FACS values internally would help to create more precise expressions and could be helpful, e.g. to train a converter from FACS to an emotional model.

A feature not yet implemented, is the selection of different face designs or even parameterizable face designs. For example: Parameterizing the level of detail of the face could prove useful if the software will run on multiple robots with differing capabilities. The level of detail could range from the simple style implemented in the prototype to a more realistic looking face for robots with close to human-like capabilities. Parameterizing the gender, skin color as well as proportions and shape of facial features may help to create a trustworthy and sympathetic faces for users with varying cultural backgrounds.

References

- [1] Kosuge, K., & Hirata, Y. (2004). Human-robot interaction. *Proceedings of 2004 IEEE International Conference on Robotics and Biomimetics* (pp. 8–11). IEEE.
- [2] Ho, C.-C., MacDorman, K. F., & Pramono, Z. A. D. D. (2008) Human emotion and the uncanny valley: A GLM, MDS, and Isomap analysis of robot video ratings. *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction HRI '08* (pp. 169–176). New York, NY, USA.
- [3] Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3-4), 143–166.
- [4] Ghent, J. (2011). *Facial Expression Analysis: A Computational Model of Facial Expression (neue Ausg.)*, Saarbrücken: VDM Verlag Dr. Müller.
- [5] Oh, J.-H., Hanson, D., Kim, W.-S., Han, Y., Kim, J.-Y., & Park, I.-W. (2006). Design of android type humanoid Robot Albert HUBO. *Proceedings of 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1428–1433). IEEE.
- [6] Guizzo, E. & Ackerman, E. (2012). The rise of the robot worker. *IEEE Spectrum*, 49(10), 34–41.
- [7] Marks, P. (2014). Robot head disguises itself as your friends. *New Scientist*, 221(2962), 20.
- [8] Kuratate, T., Matsusaka, Y., Pierce, B., & Cheng, G. (2011). Mask-bot: A life-size robot head using talking head animation for human-robot communication. *Proceedings of 2011 11th IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (pp. 99-104).
- [9] Blow, M., Dautenhahn, K., Appleby, A., Nehaniv, C. L., & Lee, D. (2006). The art of designing robot faces. *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction* (pp. 331–332). Salt Lake City, Utah, USA.
- [10] Broadbent, E., Kumar, V., Li, X., *et al.* (2013). Robots with display screens: A robot with a more humanlike face display is perceived to have more mind and a better personality. *PLoS ONE*, 8(8), 1–9.
- [11] Mori, M., MacDorman, K., & Kageki, N. (2012). The uncanny valley. *IEEE Robot. Automation Magazine*, 19(2), 98–100.
- [12] Russell, J. A. (1989). Measures of emotion. *The Measurement of Emotions — Emotion: Theory, Research, and Experience*, 4, 83–111.



Steffen Wittig was born in 1987, Tübingen, Germany. He studied media and communication informatics at Reutlingen University. He works as a software developer and graphics designer, specialized in interactive 3D and web applications. His research interests are in the fields of face animation and emotion models.



Uwe Kloos was born in 1962, Saarlouis, Germany. He studied biology (neuroscience) at Saarbrücken and Tübingen University and did his PhD in computer science at Tübingen University. He is a professor for computer graphics, visualization and collaborative virtual environments at the Reutlingen University. His research interests include working with and within virtual environments. One of his projects is the VRlab, a student's research lab focusing on VR technologies.



Matthias Rättsch was born in 1966, Neubrandenburg, Germany. He is a professor at the Reutlingen University for image understanding and interactive mobile robotics. His research interests are in the fields of personalized assistant and service systems, human robot collaboration and the multimodal interaction between humans and technical systems. He is the head of the RoboCup team RT-Lions.